

# Resonant plasmon-phonon coupling and its role in magneto-thermoelectricity in bismuth.

P. Chudzinski

Institute for Theoretical Physics, Center for Extreme Matter and Emergent Phenomena, Utrecht University, Leuvenlaan 4, 3584 CE Utrecht, The Netherlands

December 11, 2015

**Abstract.** Using diagrammatic methods we derive an effective interaction between a low energy collective movement of fermionic liquid (acoustic plasmon) and acoustic phonon. We show that the coupling between the plasmon and the lattice has a very non-trivial, resonant structure. When real and imaginary parts of the acoustic plasmon's velocity are of the same order as the phonon's velocity, the resonance qualitatively changes the nature of phonon-drag. In the following we study how magneto-thermoelectric properties are affected. Our result suggests that the novel mechanism of energy transfer between electron liquid and crystal lattice can be behind the huge Nernst effect in bismuth.

**PACS.** XX.XX.XX No PACS code given

## 1 Introduction

The thermoelectric signal, an electric response of a material upon applying a temperature gradient, encodes how entropy is transferred into the electronic system. The discovery of these highly non-trivial transport coefficients dates back to mid-nineteenth century, but recently this field enjoys renewed interest of a broad scientific community. In fundamental research it is a valuable tool to probe electron-hole correlation[1] or formation of lattice symmetry breaking orders[2]. Its role can be appreciated by a sole fact that the ultra-cold atom systems were used to build quantum emulators of this effect[3]. In the field of applied physics, there is an increasing demand for better handling with heat[4, 5], a by-product in any electric system. Devices that work as solar generators were built[6], efforts to increase figure of merit are made using nano-engineering[7], effects involving spin degree of freedom were put forward[8], and this is to name only some remarkable works from last few years. Still fundamental theoretical understanding of an origin of finite thermoelectric signal is frequently unsatisfying.

One case, where source of thermoelectricity is intensively studied[9], is the Hubbard model, where strong on-site correlations can produce Kondo features in low energy spectrum. An accepted scenario is that by lowering the temperature one strongly affects the spectral function of electrons and this in turn gives rise to a significant thermoelectric signal. Although the best known, the Kondo-physics is not the only situation, when collective phenomena strongly affect the low energy spectral function of fermions.

In this work we wish to investigate an influence of non-local correlations, collective plasmon excitations on the thermoelectric transport coefficients, which provides a very different perspective. In particular we focus on the problem how a low energy collective mode (when exist) affects energy transfer to a crystal lattice. In other words, we wish to understand better the influence of electron-electron interactions on the phonon-drag effect. For many decades the phonon-drag mechanism was successfully used and allowed researches to explain thermoelectricity in many materials, even despite the fact that theoretical understanding of mechanism as such is usually limited to single particle picture with an adiabatic approximation used to describe the coupling.

Bismuth, a material chosen as a good example to apply our theory, is a semimetal with an extremely anisotropic Fermi surface. Tiny Fermi pockets originate only because of a lattice distortion from cubic to rhombohedral. This suggest that coupling with lattice is particularly strong[10, 11]. The extreme anisotropy of Fermi surface fulfills conditions for appearance of a low energy collective excitation. Since the acoustic plasmon dominates low energy dielectric function it has to influence the low temperature phonon drag. The drag effects usually manifest in thermoelectric effects and indeed with this respect bismuth is quite intriguing. The Nernst signal in Bi is one of the largest known in nature and, as explained in Sec.A, it is beyond our understanding[12, 13, 14].

The paper is organized as follows. In Sec.2 we introduce the band structure of bismuth and briefly present the effective low energy theory. Next we give a diagrammatic derivation of plasmon-phonon coupling in Sec.3. It is shown that the the low energy plasmons can couple

with acoustic phonons and the coupling vertex has a singular structure. Next we show that the novel plasmon-phonon coupling mechanism does manifest in magneto-thermoelectric phenomena and we focus on the Nernst signal. Using results of Sec.3 we estimate (in Sec.4.2) a new phonon-drag component. Finally in Sec.5 we discuss the experimental relevance of our findings and indicate that several discrepancies in the theoretical description of thermoelectricity in bismuth can now be resolved.

## 2 Model

### 2.1 Fermi liquid: results of band theory

Bismuth is a semimetal whose Fermi surface consist out of three electron pockets (at L points of Brillouin zone) and one hole pocket (at T point). The pockets are tiny,  $k_F$  are  $10^3$  times smaller than the  $T - L$  distance[15]. The finite Fermi surface is due to a small (smaller than  $3^\circ$ ) lattice distortion which turns a cubic lattice into a rhombohedral lattice, space group A7. This implies that the small density of fermions available at lowest energies should be strongly coupled with the lattice distortions.

The hamiltonian of those carriers close to the Fermi energy reads:

$$H_{\text{fer}} = H_0^h + \sum_{i=1}^3 H_0^{e_i} + H_{\text{int}}^{h-h} + H_{\text{int}}^{e-e} + H_{\text{int}}^{e-h} \quad (1)$$

In the above  $h$  denotes a hole pocket, while there are three electron pockets indicated by  $e_i$ . The first two terms are the free fermion kinetic energies, for instance:  $H_0^h = \sum_{\mathbf{k}} E_h(\mathbf{k})(c_{\mathbf{k}}^{(h)})^\dagger c_{\mathbf{k}}^{(h)}$ , with  $(c_{\mathbf{k}}^{(\alpha)})^\dagger$  a creation operator of a fermion with momentum  $\mathbf{k}$  in the  $\alpha$  pocket. We approximate the kinetic energy  $E_\alpha(\mathbf{k})$  of each type of carrier by a free fermion dispersion relation. The peculiarity of bismuth are quite small values of certain effective masses (of order  $10^{-3}$ ) and large anisotropy of the Fermi surface ellipsoids. An effective mass ratio (in different directions) can be as large as  $10^2$  and between heavy holes and light electrons even  $10^3$  (a detailed values can be found in Ref.15). Extremely large ratios of masses of different carriers are present for most orientations of a Bi crystal. When writing Eq.1 we focus on a parabolic part of fermionic dispersion, excluding a non-parabolicity of the bands which enters into the problem for energies above 20meV. This simplification is justified as we work in the lowest temperatures, in a low magnetic field (in Bi magnetic field can in principle cause inter-band transitions) and in the long-wavelength limit where our interest is in a coupling with acoustic phonons (even high energy optical phonons in Bi have energies  $\approx 12\text{meV}$ ). Then the single particle excitations from the valence electron pockets (present at L points, at  $\approx 40\text{meV}$  below  $E_F$ ) are negligible. Nevertheless, one must keep in mind that the excluded high energy fermions contribute indirectly, for instance through a short-wavelength screening or a finite lattice elasticity that determines phonon velocity or by contributing to the

anomalous (orbital) diamagnetism of bismuth[16]. In our framework, within renormalization group spirit, the effective low energy theory emerges as a result of downfolding (integrating out) higher energy degrees of freedom, these quantities are taken as a constant parameters. In a similar way, the non-parabolic dispersion may also strengthen the acoustic plasmon (see below in Sec.2.2).

In hamiltonian Eq.1 we have also introduced hole-hole, electron-electron and electron-hole interactions. Because the Fermi pockets are so tiny (dilute gas limit) the Coulomb potential  $V_{\text{Coul}}(q)$  should be taken as a starting point. As discussed in Ref.17 a good approximation to capture the many-body effects in bismuth is the Random Phase Approximation (RPA).

### 2.2 Fermi liquid: results of RPA

The detail theory based on RPA applied to Bi was derived elsewhere[17]. Here let us only recall that the effective interaction between carriers  $V_{\text{eff}}(q)$  defined as  $V_{\text{eff}}(q) = V_{\text{Coul}}(q)/\epsilon(q, \omega)$  where  $\epsilon(q, \omega)$  is a dielectric function  $\epsilon(q, \omega) = 1 + (\Pi^e(q, \omega) + \Pi^h(q, \omega))$  obtained from polarizabilities  $\Pi^i(q, \omega)$  of fermion liquid components. The effective interaction has a resonant structure with two plasmon poles, optical and an acoustic branch. A well known optical plasmon has energy of order 10 – 20 meV, while for the other pole  $\omega \rightarrow 0$ , so it enters to the lowest energy physics.

The fact that such collective excitation may exist in a two component Fermi liquid is known already for decades[18]. The key obstacle is that the acoustic pole, fulfilling the condition  $\lim_{q \rightarrow 0} \epsilon(q, \omega = c_{\text{pl}} q) = 0$ , can be weakly damped (not overdamped) only for the material with a very anisotropic mass tensor. This condition is perfectly fulfilled in bismuth. Moreover, even if we include the non-parabolic character of  $E_e(\mathbf{k})$ , this shall amount to include an extra term  $\sim q$  in the denominator of  $\Pi^e(q, \omega)$  thus effectively shifting the effective electron mass to even smaller values. This does not change qualitatively the picture and shall even make the acoustic solution more stable. The mode has a zero sound character, but with an important difference, it has a finite width for a finite momenta. The mode is present because interaction between slower carriers are effectively screened by the rapid ones, however its finite width implies that screening is not complete (thus our model is distinct from the Hubbard model). The relation between the damping and the mass ratio can be understood if one realizes that acoustic plasmons can survive the Landau damping as long as the linear plasmon dispersion (determined by the velocity  $c_{\text{pl}}$ , see Eq.3) is above the parabola of continuum of electron-hole excitations of slower carriers. This condition gives us an upper momentum threshold  $k^*$  for the plasmons dominated regime (and, naturally, also energy  $E^* = c_{\text{pl}} k^*$  and temperature  $E^* = k_B T^*$  thresholds).

In this regime the hamiltonian Eq.1 can be reduced to an effective low energy hamiltonian  $\tilde{H}_{\text{fer}}$  which consist of free electrons and holes coupled to the collective bosonic

excitation:

$$\tilde{H}_{\text{fer}} = H_0^h + H_0^e + \sum_q \omega_q b_q^\dagger b_q + \frac{1}{\sqrt{\Omega}} \sum_{\alpha=e,h} \sum_q M_q^\alpha [b_{-q}^\dagger + b_q] \rho_\alpha(q) \quad (2)$$

where  $b_q^\dagger$  is a boson (plasmon) creation operator and:

$$\begin{aligned} \omega_q &= c_{\text{pl}} |q| \\ M_q^{e0} &= \left( \frac{3\pi^2 m_r c_{\text{pl}}^3}{2k_{Fr}^3} \right)^{1/2} |q|^{1/2} \\ c_{\text{pl}} &= \sqrt{V_{Fr} V_{Fs} / 3} \end{aligned} \quad (3)$$

where  $V_{Fr(s)}$  is the Fermi velocity of rapid (slow) carriers,  $k_{Fr(s)}$ ,  $m_r(s)$  are their respective Fermi wavelengths and effective masses, and  $\rho_\alpha(q) = \sum_k c_{k+q}^\dagger c_k$  is a density of carriers from  $\alpha$  pocket. The  $M_q^{e0}$  is a  $q$  dependent probability of a boson decay back into an e-h pair (this determines plasmon damping, or inverse lifetime, in the electronic system). In our problem both slow and rapid carriers has a finite density at  $E_F$  thus we must keep both in Eq.2.

### 2.3 Lattice: coupling with collective mode

We proceed by incorporating lattice into our considerations. Then the total hamiltonian reads:

$$H_{\text{tot}} = \tilde{H}_{\text{fer}} + H_{\text{ph}} + H_{\text{coupl}} \quad (4)$$

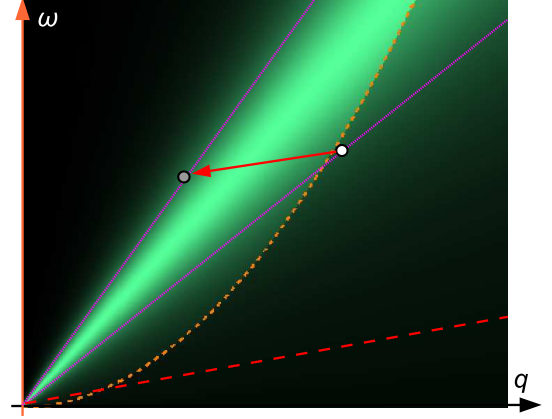
In the above equation  $\tilde{H}_{\text{fer}}$  states for Fermi-liquid hamiltonian as defined in (2). The lattice hamiltonian  $H_{\text{ph}}$  is an acoustic phonon bath with a linear spectrum:

$$H_{\text{bath}} = \sum_k c_{phk} a_k^\dagger a_k \quad (5)$$

where  $a_k^\dagger$  creates a phonon with momentum  $k$ .

$H_{\text{coupl}}$  is, yet unknown, coupling between acoustic phonon and acoustic plasmon. A standard coupling, through a displacement potential, a change of a single electron eigenenergies upon lattice distortion (thus a direct coupling with the  $H_0$  part of Eq.1), is assumed to be known and not of interest in this study.

Another effect, that is very pronounced in bismuth, is that a lattice distortion changes the volume of electron/hole pockets (this is allowed in a semimetal). The effect is the strongest for the band bottom, the parabolic part of  $E(k)$ . When  $k \neq 0$ , then modulation is not uniform in space, carriers have to move to accommodate within the new energy landscape. Again, an extreme ratio of light/heavy carrier masses, allows for a new phenomenon to occur. Different accelerations induce a finite polarization of the two component Fermi liquid and this locally produces a finite electric field  $E_{\text{ind}}$ . Since the induced electric field potential  $\nabla V_{\text{ind}} = E_{\text{ind}}$  is proportional to lattice distortion, thus such coupling can be interpreted as an effective piezo-electric interaction[19]. It is known[19] that



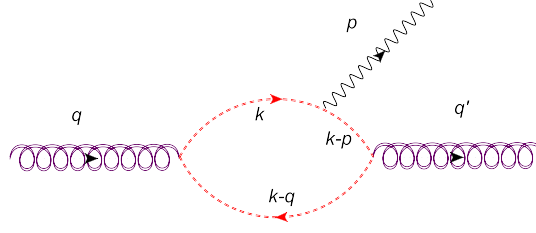
**Fig. 1.** Phonon emission process (red vector) on energy-momentum plane. The green shading (bordered by two purple guide-lines) indicates weight of plasmon spectral function. Initial plasmon state (white dot) lies on the lower edge of the spectral function, while the final (grey dot) on the upper edge. Under the orange dashed line (a parabola) there is a zone where strong Landau damping by slow fermions sets in. Red dashed line shows phonon dispersion, it is parallel to the red vector. In the illustrated case one achieves  $q_{ph} \approx q_{pl}$ .

$V_{\text{ind}}$ , which is phase shifted by  $\pi/2$ , does not interfere (at least up to a second order) with effects of displacement potential.

Physically, an additional contribution arises thanks to carriers' collective motion – numerous fermionic states buried deeply below  $E_F$  are able to couple with low energy lattice dynamics. Clearly, there is no double counting: different groups of fermions drag phonons via the standard mechanism or via the plasmonic channel. A possibility of direct coupling between phonons and plasmons have been indicated already before, by a numerical evaluation of integrals for  $T = 0$  charge RPA susceptibilities in low dimensional structures [20, 21]. In the following section we prove that a relevant scattering amplitudes are indeed non-zero and evaluate their momentum dependence.

### 3 Plasmon-phonon coupling

Existence of the tiny Fermi pockets induced by distortion is one crucial property of bismuth. Below we derive the resonant coupling which reveals yet another prerequisite. As illustrated in figure Fig.1, it is the existence of plasmonic and bosonic dispersions that fall sufficiently close on the momentum-energy plane. In a particular case of linear dispersion, it is a finite width of a plasmon that allows to compensate the difference of the two velocities. In bismuth, the desired acoustic plasmon exist, thanks to highly anisotropic tensors of effective masses for both electrons and holes.



**Fig. 2.** Emission of a phonon with momentum  $p$  (black zig-zag line) by a plasmon with momentum  $q$  (purple double solenoid line) with simultaneous emission of a plasmon with momentum  $q'$ . Red dashed lines are fermions' propagators inside an extracted bubble. This three leg bubble diagram represents the third order polarizability  $\Pi^{III}(\omega, q, q')$ .

### 3.1 Bare plasmon-phonon vertex

Bismuth crystal structure possesses an inversion symmetry, thus only the dynamically induced piezo-electric coupling is allowed. Then an issue is how a polarizability of Fermi liquid changes upon emission/absorption of a phonon. When the polarizability of the considered Fermi liquid has a singularity for certain  $\omega, k$  then it should dominate the response. It is then justified to use the plasmon-pole approximation.

The simplest process is a direct transformation of a propagating plasmon into phonon. The velocities of plasmons and phonons differ by a factor four, thus energy-momentum conservation prevents such events. In order to conserve energy and momenta the second plasmon must be simultaneously emitted as it is shown in Fig.1. Then the processes to be considered is plasmon destruction with simultaneous emission of another plasmon and phonon which can be also seen as a coherent plasmon backscattering on the lattice. This is shown on Fig.2 and described by the following hamiltonian:

$$H_p = \sum_{q'} \Upsilon(q, q') b_q a_{q-q'}^\dagger b_{q'}^\dagger + \text{h.c.} \quad (6)$$

In a diagrammatic language, to evaluate  $\Upsilon(q, q')$ , we extract a fermionic bubble from the acoustic plasmon propagator<sup>1</sup>. This bubble can emit a phonon with the momentum  $q_{ph} = q - q'$  and a plasmon with momentum  $q_{pl} = q'$ . The emission of the plasmon, with properly tuned energy and momentum is possible exclusively due to a finite life-time of acoustic plasmon as illustrated on Fig.1.

Then the  $\Upsilon(q, q')$  is:

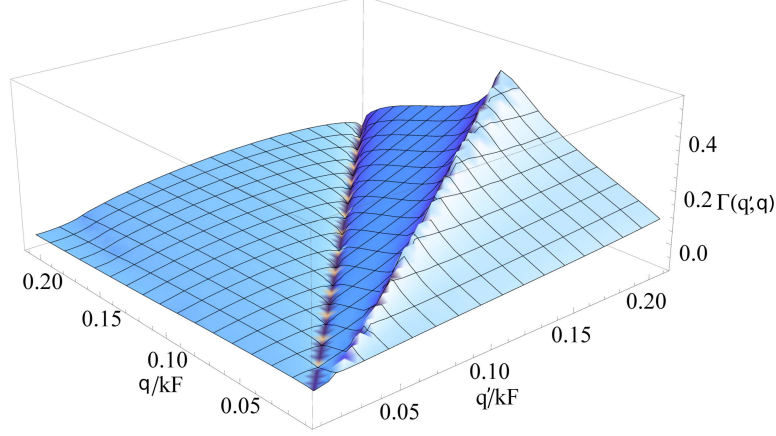
$$\Upsilon(q, q') = M(q) V_{\text{ind}}(q - q') M(q') \Pi^{III}(\omega, q, q')|_{\omega=0.8c_{pl}q} \quad (7)$$

where the first term  $M(q) \sim \sqrt{q}$  accounts for a plasmon conversion into a bubble (particle-hole pair) and the second term  $M(q')$  is a conversion of a bubble back into plasmon,  $V_{\text{ind}}(q - q')$  is the induced piezo-electric potential which acts on a lattice,  $\Pi^{III}(\omega, q, q')$  is a higher order polarizability of plasma. The  $\Pi^{III}(\omega, q, q')$  is a three leg bubble diagram: three bosonic propagators leave a single bubble and one of these lines is distinguishable from the others, see Fig.2. We pick the  $\omega$  value based on quantitative analysis given in Fig.1, but we have checked that the precise value does not affect the result, the resonant structure is always present. The three leg bubble diagram (drawn on Fig.2) can be expressed as follows:

$$\Pi^{III}(\omega_n, \mathbf{q}, \mathbf{q}', \Omega_2^{ph}) = \sum_{\omega_1, k} G(k - q, \omega_1 - \omega_n) G(k, \omega_1) G(k - p, \omega_1 - \Omega_2^{ph})|_{p=q-q'} \quad (8)$$

The finite temperature free fermion propagator is taken as  $G(k, \omega_1) = (\omega_1 - \epsilon_k)^{-1}$  (with  $\omega_1$  being fermionic Matsubara frequency) where  $\epsilon_k$  is a single particle eigen-energy of the heavier carriers (it either  $E_h(\mathbf{k})$  or  $E_e(\mathbf{k})$ , in both cases an even function of  $\mathbf{k}$ ). In order to evaluate  $\Pi^{III}(\omega, q, q', \Omega_2^{ph})$  we proceed in a way very similar to derivation of the (two leg) Lindhard function (see App.D for details). As a result at zero temperature we get the following formula (see App.D for derivation):

<sup>1</sup> Since the bare Coulomb line goes into an optical plasmon branch, an acoustic plasmon propagator always has a fermionic bubble inside its propagator. One extracts a bubble out of an infinite series from which the phonon can be emitted (see Fig.2). Upon this operation the RPA series remains re-summable and the plasmon propagator is still well defined, in fact it can even increase its spectral weight thanks to translational symmetry breaking by the lattice distortion.



**Fig. 3.** A real part of the scattering vertex  $\Upsilon(q, q')$  on the  $q'$ - $q$  plane: an incoming plasmon with momentum  $q$  scatters into another plasmon with momentum  $q_{pl} = q'$  and emits a phonon with momentum  $q_{ph} = q - q'$ .

$$\begin{aligned} \Pi^{III}(\omega, \mathbf{q}, \mathbf{q}', \Omega_2^{ph} \rightarrow 0) = \int_0^{k_F} d^3\mathbf{k} \left( \frac{1}{(\omega - \epsilon_{\mathbf{k}+\mathbf{q}} + \epsilon_{\mathbf{k}} + i\delta)(\omega - \epsilon_{\mathbf{k}} + \epsilon_{\mathbf{k}-(\mathbf{p}-\mathbf{q})} + i\delta)} + \right. \\ \left. \frac{1}{(\omega - \epsilon_{\mathbf{k}} + \epsilon_{\mathbf{k}-\mathbf{q}} + i\delta)(-\epsilon_{\mathbf{k}-\mathbf{p}} + \epsilon_{\mathbf{k}} + i\delta)} + \frac{1}{(\omega + \epsilon_{\mathbf{k}+(\mathbf{p}-\mathbf{q})} - \epsilon_{\mathbf{k}} + i\delta)(\epsilon_{\mathbf{k}+\mathbf{p}} - \epsilon_{\mathbf{k}} + i\delta)} \right) \Big|_{p=q-q'}, \end{aligned} \quad (9)$$

The momentum integral can be, in some cases, performed analytically but formulas are rather long and thus given in the appendix (App.D). In the main text we only present the result, figure Fig.3, obtained for the specific set of parameters:  $Re[V_{pl}] = 4Im[V_{pl}] = 4V_{ph}$ , which represents a particular case of bismuth.

Figure Fig.3 reveals a highly non trivial structure of the scattering vertex  $\Upsilon(q, q')$ . There is a strong resonance in momentum dependence spanned around the line  $q' \approx q/2$  which is equivalent to  $q_{ph} = q_{pl}$ .  $Re[\Upsilon(q, q/2)]$  approaches values of order  $10^{-1}$ , which on one hand implies that our perturbative approach is reasonable, on the other that plasmon-phonon coupling have a non-negligible strength.

Physically, existence of the resonance becomes clearer if one considers an opposite process: when combined plasmon and phonon produces a plasmon wave. This can be interpreted as a constructive interference of two displacement waves. One then indeed expects a resonance around  $q_{ph} = q_{pl}$ . Importantly, the resonance condition restricts not only the value of wavevectors, but also their direction,  $\mathbf{q}_{ph} \parallel \mathbf{q}_{pl}$ .

In Fig.3 one easily notices that strength of  $\Upsilon(q, q/2)$  increases with  $q$ . The momentum dependence of the resonant part is crucial for our further reasoning in Sec.4.2, so we investigate it in detail. First we check, on log-log plots that  $\Upsilon(q, q/2)$  can be approximated quite well by a power law, at least when we restrict ourselves to  $q < 0.1k_F$  ( $q^* < 0.1k_F$ ). For further calculations we take:

$$\Upsilon(q, q/2) = \bar{\Upsilon} q^\alpha \quad (10)$$

The exponent varies (slightly) only upon changing the ratio  $c_{ph}/c_{pl}$ : it is reduced from 0.57 for  $c_{ph}/c_{pl} = 1/4$  down to 0.52 for  $c_{ph}/c_{pl} = 1/2$ , while changing other parameters even by factor five does not affect it. At the same time the amplitude of  $\Upsilon(q^*, q^*/2)$  may vary a few times. Although the precise amplitude the coupling does depend on fine details of the model, the exponent  $\alpha \approx 0.5$  is universal enough to allow for detection when fitting experimental results, which are routinely checked for the power-law behavior.

## 4 Nernst signal in bismuth

A question is whether a plasmon-phonon resonance that we have found in previous section is detectable. A good candidate, where plasmon mediated phonon drag can manifest, are the thermoelectric phenomena.

### 4.1 Kelvin-Thomson relation

We compute only a temperature dependence of the Nernst signal that is characteristic for the novel mechanism. A straightforward 'S'-approach, within the Nernst configuration, is rather unpleasant from a theoretical viewpoint: one must introduce variations of temperature within the sample, which is a fundamental issue since the heat current will also contain variations of chemical potential. Fortunately thanks to the Onsager relation (known as Kelvin-Thomson relation in this case) we can turn to a reciprocal coefficient, the Peltier coefficient  $\Pi_{xy}$ . Then the Nernst signal  $N \equiv S_{xy}$  can be obtained from:

$$S_{xy}(T; B) = \Pi_{xy}(T; -B)/T \quad (11)$$

from Eq.11 we see that Nernst signal can be associated with entropy transport in the system. Now one compute the Nernst coefficient using so called 'II'-approach, in the Peltier configuration where a heat transfer is induced by an electric voltage. Precisely, we will calculate a heat current  $\dot{Q}_y$  induced by a flow of electric current  $j_x$  (under a perpendicular magnetic field  $B_z$ ). By definition the Peltier coefficient  $\Pi_{xy}$  is the proportionality constant:  $\dot{Q}_y = \Pi_{xy} j_x$  where  $\dot{Q}$  is the time derivative of heat diffused in y-direction. Thus, the price we need to pay in the 'II'-approach is that we have currents flowing, a non-equilibrium situation, in two perpendicular directions.

### 4.2 Plasmonic contribution

First compute an amount of heat transported by plasmons, or plasmonic clouds, accompanying fermions. As suggested by experimental findings in bismuth[22] we assume that in its movement each electron is accompanied by a finite number of plasmons  $n_{pl}$ , a so called plasmonic picture. By analogy with polarons[23] we know that in average  $\langle n_{pl} \rangle = M_q$ . We imposed a constant electric current  $j_x$ , which can be associated with a voltage difference  $V_x = \rho_{xx} j_x$ . In the presence of crossed fields  $E_x$ ,  $B_z$  the electrons and holes move in the opposite directions along the  $x$ -axis, but due to Lorentz force towards the same electrode in  $y$ -direction. This is the  $E \times B$  drift, with constant velocity  $u_y = E_x B_z$ . It brings an extra entropy (heat) to one side of conducting slab.

Then the heat transfer in  $y$  direction is:

$$\dot{Q}_y = j_x \rho_{xx} B_z \int dq^3 c_{pl} q M_q b(\beta c_{pl} q) \quad (12)$$

with an integral which is very much like in a Debye model for specific heat, except an extra  $M_q$  term. As

usual we extract temperature dependence by rescaling the momentum  $\tilde{q} = \beta c_{pl} q$  (to get dimensionless argument of Bose-Einstein distribution). Following these steps we find:

$$\Pi_{xy}(T; -B) \sim \rho_{xx} c_{pl} \bar{M} B_z T^{3.5} \quad (13)$$

where  $\bar{M} = M_q / \sqrt{q}$  is momentum independent. In passing we note that  $\Pi_{xy} \sim \rho_{xx}$  and this, for the case of Bi where  $\rho_{xx} \gg \rho_{xy}$ , is in agreement with a general tensorial relation between transport coefficients  $\hat{\Pi} = \hat{\alpha} \cdot \hat{\sigma}^{-1}$ , where  $\hat{\alpha}$ ,  $\hat{\sigma}$  are thermoelectric and electric transport coefficients respectively. From Eq.13 and Eq.11 it immediately follows that  $S_{xy} \sim T^{2.5}$ .

In bismuth  $c_{pl}$  is only slightly larger than  $V_{Fs}$  (Fermi velocity of slower carriers) and  $\bar{M}$  is small. This implies that an amplitude of plasmonic contribution to Nernst signal, Eq.13 cannot be much larger than that of the Fermi liquid. Then, in the lowest temperatures, a sole plasmonic contribution ( $\sim T^{2.5}$ ) cannot dominate the Fermi liquid contribution ( $\sim T^1$ ).

### 4.3 Phononic contribution

When the flow of fermion liquid induces lattice displacement the main contribution to thermo-magnetic heat transfer will come from the large heat capacity of the phonon gas. This is a realization of Gurievich mechanism[24] in the presence of collective excitations. We follow an approach frequently used for drag phenomena, see Ref.[25] for a specific case of bismuth, and express  $\Pi_{xy}$  as an extra heat that is transferred to the lattice:

$$\dot{Q}_y = \int \frac{d^3 q}{(2\pi)^3} c_{ph} q \cdot \hat{y} g(\mathbf{q}) \quad (14)$$

where  $\hat{y}$  is a versor in  $y$ -direction and each phonon brings (or takes) an energy  $c_{ph} q$ . The  $g(\mathbf{q})$  is a displacement of a phonon distribution  $N_q$  from its equilibrium Bose distribution  $N_q^0$ :  $g(\mathbf{q}) = N_q - N_q^0$  due to applied fields. The  $g(\mathbf{q})$  can be found from the Boltzman equation for the steady state of phonons:

$$\left( \frac{\partial N_q}{\partial t} \right)_{\text{force}} + \left( \frac{\partial N_q}{\partial t} \right)_{\text{relax}} = 0 \quad (15)$$

where the two terms, that are balanced in the steady state, are due to external forces e.g. due to  $\mathbf{E}$ ,  $\mathbf{B}$  and relaxation e.g. due to boundaries, impurities. Formally, Eq.15 should be supplemented with three more equations, for plasmons and slow/fast fermions. Such system of partial differential equations could be treated only numerically. However, a few basic properties of bismuth, outlined in App.B, allowed us to restrict to Eq.15 and obtain an analytic expression.

We take  $\frac{\partial N_q}{\partial t}_{\text{relax}} = -g(\mathbf{q})/\tau_r$ . This is justified for the phonons whose distribution can relax mostly by scattering on lattice imperfections (like for instance twin boundaries known to be present in a rhombohedral lattice).

The first term in Eq.15, due to the effective fermion-phonon scattering, is estimated in Born approximation:

$$\left(\frac{\partial N_q}{\partial t}\right)_{\text{force}} = \sum_{n,n'} [W^{em}(n', n) f_{n'}(1 - f_n) - W^{ab}(n', n) f_n(1 - f_{n'})] \quad (16)$$

where  $f_{n'}$  is a shorthand notation for a Fermi-Dirac distribution at energy  $E_{n'}$ . These energies does depend on external forces, the fields  $E_x$  and  $B_z$ , thus one can Taylor expand Eq.16 up to  $O(E_x)$ . The procedure is standard and we refer to original paper Ref.26. The Taylor expansion brings a factor  $\partial f(E)/\partial E$  which in the following (e.g. in Eq.18) limits integration to a 2D Fermi surface (in all our considerations  $k_B T \ll E_F$ ). The novel two stage scattering enters through transition probabilities of phonon absorption/emission

$$W^{ab}(n', n) = b(\beta c_{ph} q) W(n', n);$$

$$W^{em}(n', n) = (b(\beta c_{ph} q) + 1) W(n', n)$$

where the bare probability  $W(n', n)$  is evaluated from the Kramers-Heisenberg formula:

$$W(n', n) = \left( \frac{M_q \Im(q, q/2)}{E(k+q) - E(k) - \text{Im}[c_{pl}]q} \right)^2 \quad (17)$$

where the difference between initial  $n$  and final state  $n'$  is a phonon with momentum  $q/2$ . The intermediate state of the process, that is able to emit a phonon, is a plasmaron. From Eq.2 a probability of creating an intermediate state, a plasmon that accompanies fermion, is  $M_q$ . A probability that a phonon will be emitted, from Eq.6 is  $\Im(q, q/2)$ . The last formula, Eq.17, closes our system of equations and by using Eq.16 together with Eq.17 we can solve Eq.15 for  $g(q)$ . This we can substitute back into Eq.14 to find  $\dot{Q}_y$ , or after dividing by  $j_x = E_x/\rho_{xx}$  to find directly the Peltier coefficient:

$$\Pi_{xy}(T) = \rho_{xx} \int_0^{q^*} d^2 q c_{ph} q W(n, n') b(c_{ph} q/2) \quad (18)$$

Comparing Eq.14 with Eq.18 we see that

$$W(n, n') b(V_{ph} q/2) = g(q)$$

while from Eq.15 we know that  $g(q) \sim \tau_p^{-1}$  where  $\tau_p^{-1}$  is phonon scattering rate due to interaction with plasmons. By comparison with results of App.C we see that the term  $(M_q \Im(q, q/2))^2 b(c_{ph} q/2)$  in Eq.18 is equal to the imaginary part of the phonon self-energy (Eq.27) times the plasmon spectral weight. The link between Eq.18 and Eq.27 can be interpreted as a generalized version of a Fermi golden rule with an inverse scattering time  $\tau_p^{-1}$  equal to  $\text{Im}[\Sigma_p(q, \omega)]$ .

We now use the fact that the denominator of Eq.17 scales linearly with momentum, while momentum dependence of  $M_q \sim \sqrt{q}$  and  $\Im(q, q/2) \sim q^\alpha$  to find  $W(n, n') \sim$

$q^{2\alpha-1}$ . Being primarily interested in the temperature dependence at the lowest energies we re-scale momentum  $\tilde{q} \rightarrow \beta q$  (the only temperature dependence in Eq.19 comes from an argument of the Bose-Einstein distribution  $\sim \beta q$ ). We find:

$$\Pi_{xy}(T) = T^{2(1+\alpha)} \rho_{xx} A \int_0^{\tilde{q}^*} d^2 \tilde{q} c_{ph} \tilde{q} \tilde{q}^{2\alpha-1} b(c_{ph} \tilde{q}/2) \quad (19)$$

where a momentum independent constant

$$A = \left( \frac{3\pi^2 m_r c_{pl}^3}{2k_{Fr}^3} \right)^{1/2} \Im/(c_{ph} - \text{Im}[c_{pl}])$$

The integral is known:

$$\int_0^{\tilde{q}^*} d^2 \tilde{q} c_{ph} \tilde{q} b(c_{ph} \tilde{q}/2) = Li_2(\exp(c_{ph} \tilde{q}^*)) + c_{ph} \tilde{q}^* \log(1 - \exp(c_{ph} \tilde{q}^*)) - (c_{ph} \tilde{q}^*)^2/2 \quad (20)$$

where  $Li_2$  is a di-logarithm function. In the limit of zero temperature  $\tilde{q}^* \rightarrow \infty$  and the right hand side of Eq.20 becomes temperature independent. Then, from Eq.11,  $N(T) = \Pi_{xy}(T)/T$  and we obtain the following temperature dependence for the Nernst signal:

$$N(T) \sim T^{1+2\alpha} \quad (21)$$

On the other hand for higher temperatures, when  $k_B T \geq c_{ph} q^*$ , the temperature dependence of the right hand side of Eq.20 matters. We approximate it by the last term  $\sim \tilde{q}^{*2} \equiv T^{-2}$ , to find that in high temperatures  $N(T)$  should slowly saturate to a constant value. The regime of finite boson life-time is covered in our formalism, the saturation profile will broaden and the saturation value decrease when one includes finite imaginary velocity in the argument of Eq.20. What we have neglected so far, is an intense Landau damping of plasmons that takes place above the characteristic temperature  $T^*$ . This suppresses the number plasmons available in the system. Since the plasmons are a necessary ingredient of our drag mechanism, the  $W(n, n')$  is also suppressed and the  $N(T > T^*)$  follows this decrease (see Sec.5 for further discussion).

## 5 Discussion, comparison with experiments

The predicted full temperature dependence of the Nernst signal is shown on Fig.4. In the lowest temperatures the consideration of the previous section applies and the thermoelectric signal is  $\sim T^{1+2\alpha}$ . Above the characteristic temperature  $T^*$  (see Sec.2.2) the Coulomb interaction are no longer mediated by collective excitations: fermions excited around  $E_F$  are too energetic to be followed by acoustic plasmons instead they cause strong Landau damping.

Experimentally, in bismuth[27, 12], and for the entire  $Bi_{1-x}Sb_x$  family of semimetals as well ( $x < 0.06$ )[28], a

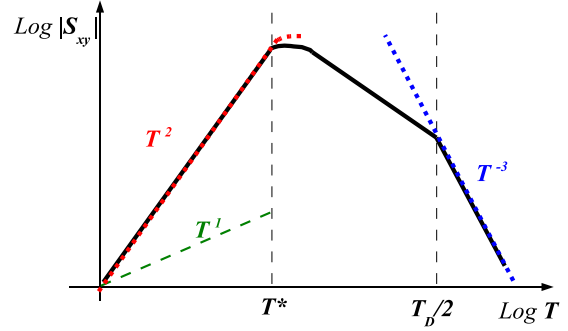


power law  $N(T) \sim T^b$  is clearly observed, with  $b \approx 2$ , for temperatures below  $T \approx 3.5K$ , a behavior that is clearly different than the one expected for the Fermi liquid. The maximum at  $3.5K$  is strikingly similar to the upper limit of plasmonic-type Baber resistivity[29] and the previous estimate[17]  $T^* \approx 2 \div 4K$ . No other characteristic energy scale of the material falls in this energy range. Above  $3.5K$  a decay of the Nernst signal was observed.

At higher temperatures, for  $T > 20K$  the Nernst effect was explained in Ref.[30]. In that work a standard single-fermion phonon drag was considered in the regime where the phonon-phonon scattering dominates relaxation. The  $N \sim T^{-3}$  decay was predicted, while the presence of temperature dependent relaxation could be detected as deviation from this power law  $N \sim T^{-3+\delta}$  where  $0 < \delta < 1$ . This was successfully compared with experimental data from Farag and Taruma[30]. The invoked amplitude of electron-phonon coupling falls closely to a recent *ab-initio* estimate[31]. The  $N \sim T^{-3}$  was found for  $T > 80K$ , while at lower temperatures  $N \sim T^{-2}$  fits better which is an indication that  $1/T$  temperature dependence of the relaxation rate is present and one needs to go beyond standard electron-phonon coupling theory. This is what is achieved in the present study.

We should note two recent theoretical studies of Nernst effect in bismuth, although these are primarily devoted to the high magnetic field regime. First work, Ref.[25], by invoking only the phonon-drag mechanism, within the "H"-approach, allowed to successfully explain the positions of peaks in the Nernst signal at high magnetic fields. Second work, Ref.[32], treated the same problem, but primarily focusing on the electronic, dissipationless contribution. It has also allowed to successfully explain the experimental signal, however a substantial phonon-drag contribution (of the same order as electronic part) needed to be accounted for. In both works the Nernst signal is proportional to longitudinal resistivity  $\rho_{xx}$  and fermionic density, in agreement with the novel mechanism postulated here. In both works strong temperature dependence was found, for instance in Ref.[25] it is certainly stronger than linear dependence. The temperature dependence of phenomenological phonon-drag postulated (at low magnetic fields) in Ref.[32] turns out to be one power faster than the dissipationless component, and falls close to  $T^2$ . This is in disagreement with our Eq.21, the main finding of our work. To complete this comparison, it should be clearly stated that in high magnetic fields the novel mechanism proposed here needs to be substantially modified: since then a magneto-plasmon has a quadratic, massive dispersion it will couple with an optical phonon, the amplitude of the lattice distortion.

Furthermore, we compare the transverse (Nernst) and longitudinal (Seebeck) signals. We use axes convention from Sec.4.2. In an ambipolar material, along the  $x$ -axis electrons and holes move towards the opposite electrodes, thus accompanying them plasmons are dragged in the opposite direction along this axis. The longitudinal, Seebeck part of thermoelectric signal is then expected to be small. Upon a presence of magnetic field, both electrons and



**Fig. 4.** The predicted temperature dependence of the Nernst signal. The  $T^2$  power law at the lowest temperatures is found from our novel plasmon-phonon coupling mechanism (red dotted line). The  $T^1$  behavior (green dashed line) would be there for the Fermi liquid case. We also indicate the result of previous phonon-drag studies in in Ref.[30] (blue dotted line): in the highest temperatures a phonon drag with a power law decay  $N \sim T^{-3}$  was predicted. For comparison an experimental result is sketched as a black solid curve. In experiments it was found that the lowest temperatures  $T^2$  growth saturates around  $3.5K$  ( $\approx T^*$ ) while a regime where  $N \sim T^{-3}$  is valid terminates around  $50-80K$  that is approximately  $\approx T_D/2$ .

holes deviates in the y-direction towards the same edge, the dragged plasmons and phonons can then build up a large difference of temperatures. This implies that, at the lowest temperatures, for plasmon-phonon drag  $S_{xx} \ll S_{xy}$  and indeed this is observed in the experiment[12]. The anomalous low temperature signal is much stronger in the transverse (Nernst) than in longitudinal (Seebeck) effect. In principle the phonon-plasmon regime is limited also on the low temperature side, by the Fermi liquid component  $\sim T$ . The amplitude of this component is much smaller, and indeed such behavior was not detected even down to mK temperature range in the transverse, Nernst signal, while competition between different mechanisms has been observed in the Seebeck signal[33].

Finally let us look at propagation of phonons. While they experience mixing with plasmons, there are no point-like collisions with fermions or defects. The drag process is driven by the normal (not umklapp) scattering with other bosons which implies that the only effect is the renormalization of phonon velocity. In this low temperature regime, in the transport experiment, the phonons can be detected as ballistic[27], but their velocity should be slightly larger than expected. It is expected that heat transfer is dominated by the bosonic bath and indeed ballistic transport was found when measuring thermal conductivity of bismuth[27]. Moreover, a sudden drop of thermal conductivity was observed at  $T \approx T^*$ , which suggest a profound link between the thermal transport and the thermo-magnetic signal.



One may also consider more sophisticated inverse probes – like the change of sound propagation upon applying magnetic field. According to the novel plasmon-phonon coupling proposed here they can be affected. Indeed anomalously large oscillatory sound attenuation (an analog of Shubnikov-de Hass) were experimentally observed [34, 35, 36] again at temperatures of a few kelvins. Detail study of these phenomena, as well as anomalous attenuation of hot electrons in Bi, can provide further evidences of the novel mechanism of coupling with a lattice.

## 6 Conclusions

In this work we have investigated influence of an acoustic plasmon on a coupling between Fermi liquid and a lattice. We found quite non-trivial momentum dependence of the coupling vertex  $\Upsilon(q, q')$  with an emergent resonance where plasmon-phonon coupling can be significant. The resonance is reachable only thanks to a finite imaginary part of the acoustic plasmon velocity. This is quite distinct from the short range Hubbard-like models where the zero-sound is infinitely sharp. Our model is dedicated to extremely diluted regime, where non-local correlation effects are dominant. One manifestation of the plasmon-phonon resonance is a novel component of the phonon drag – plasmon mediated phonon drag. Bismuth is indicated as a likely material where such coupling can take place and indeed we are able to provide an interpretation for the huge low temperature Nernst signal in this element.

There are other materials, for which our results are relevant. One prominent example are di-chalcogenides whose anisotropy of the Fermi surface falls close to the one of bismuth. An unusual charge density wave transition, that sometimes is ascribed to the formation of an excitonic insulator[37], takes place in these materials. An issue of the role and nature of electron-phonon coupling is a subject of intense ongoing research[38], electron-phonon vertices with strong momentum dependence were recently invoked[39, 40] to explain some of the peculiarities.

Our results may be useful from applied physics viewpoint, in low dimensional (1D and 2D) nano-structures. In these structures plasmons are always acoustic ( $\omega(q \rightarrow 0) \rightarrow 0$ ) and their velocity (real and imaginary part) can be to some extent tuned. Our work shows that by varying the plasmon's velocity one can significantly change the magnitude of phonon drag. Two applications are possible, the first, mentioned already in the introduction, is to optimize for a maximal magneto-thermoelectric signal, in a desired direction, by enhancing the plasmon-phonon resonance. The second, that stems from a growing interest in quasi-2D plasmonics, is to suppress the phonon-drag resonance in order to minimize losses due to non-radiative decay of acoustic plasmons.

## Acknowledgments

I would like to thank Thierry Giamarchi for many inspiring discussions, especially in the initial stage of this project.

## Author contribution statement

The author solely performed all the calculations and wrote the manuscript.

## A Thermoelectricity in Bi, a state of knowledge

The Nernst signal measures a perpendicular electric voltage caused by a temperature gradient (in the presence of magnetic field):

$$N = \hat{\sigma}^{-1} \hat{\alpha} = \frac{\alpha_{xy} \sigma_{xx} - \alpha_{xx} \sigma_{xy}}{\sigma_{xx}^2 + \sigma_{xy}^2} \quad (22)$$

The Sondheimer cancellation between Hall and transverse thermoelectric voltages, the numerator in Eq.22, means that the signal is related to electron-electron interactions (the non-interacting electrons propagator gives a zero signal[19]). For a metal behaving according to a semi-classical theory the Mott relation  $\alpha \sim \partial\sigma/\partial\omega$  is obeyed. The Mott relation express the fact that non-zero thermoelectric power  $\alpha_{ij} \sim \langle J_i; J_j^E \rangle$  can appear only from the energy (momentum) dependent part of self-energy[41]. Despite these cancellations there are attempts to ascribe the extremely large amplitude of the Nernst signal solely to a standard Fermi liquid component. In Ref.27 a very small value of the Fermi energy and a very long electron scattering time were claimed to be responsible and a following relation was given:

$$N = \frac{(\pi k_B)^2 T}{3} \frac{\partial \log(\sigma)}{\partial \omega} \approx \frac{\pi^2}{3} \frac{k_B^2 T}{m^*} \frac{\tau}{E_F} \quad (23)$$

The critical approximation done in Eq.23 are Mott relation, assumption  $\sigma_{xx} \gg \sigma_{xy}$  (and obviously the same for  $\hat{\alpha}$ ), Taylor expansion of a logarithm and furthermore  $\partial\sigma/\partial\omega = \tau^0/E_F$ , where a hidden assumption about the relaxation time  $\tau^0$  (which is supposed to approximate the electrons scattering vertex) being a constant. The low energy effective theory, Eq.2 produces fermionic self-energy that strongly depends on energy-momentum. This immediately implies that Eq.23 does not hold. Even if one decides to forget about this important component of self-energy, in a system with an extremely small  $E_F$  like Bi the band curvature matters esp. when approximating the derivative (Eq.23) by the ratio, thus Eq.23 should be thought as an upper limit for  $N$ . Strikingly,  $N$  estimated by Eq.23 is one order of magnitude smaller than the observed ones (it is clear from Fig. 2 in 27). Moreover the reasoning so far neglects one further cancellation: bismuth is a semimetal

so the mobilities of electrons and holes partially cancel, which furthermore reduces the value estimated from Eq.23.

According to Eq.23,  $N(T)$  is expected to scale linearly with temperature and this law should dominate as  $T \rightarrow 0$ . This is in a disagreement with experimental findings, a power law with larger exponent has been found, see Sec.5 for details. There is an additional puzzling element discussed in Ref.[25]. Experimentally [28, 12] one has the remarkable hierarchy  $\sigma_{xx} \gg \sigma_{xy}$  (which is typical for semimetals) and  $\alpha_{xx} \ll \alpha_{xy}$ . The two relations definitely cannot fit together when the Mott relation holds. To summarize: we do not claim that the dissipative component of thermoelectric signal is zero, but that it cannot be the dominating one at low temperatures.

The other component of thermoelectric power, not accounted in the Fermi liquid theory, comes from the phonon drag. Thus the phonon-drag processes are never accounted for in the approximation Eq.23. The Mott relation, giving the first equality in Eq.23, holds and gives all contributions to  $N$  only provided that fermions are scattered on a static potential (of any origin)[41]. Phonon-drag mechanism does not fulfill this condition.

## B Approximation of Boltzman equation

Our approximation is that fermions are strongly coupled with plasmons and relax only through plasmon-phonon scattering. This can be justified by:

- in bismuth there is  $\sim 10^4$  atoms per one electron (tiny Fermi pockets), then lattice has much larger number of degrees of freedom than electron liquid. According to second law of thermodynamics entropy will flow from fermionic liquid to phonon bath.
- in low temperatures available electronic wavelengths are four orders of magnitude larger than lattice parameters, thus scattering with point defects is negligible<sup>2</sup>, on the contrary the resonant plasmon-phonon coupling is rather strong so all the energy transfer must be done through this scattering channel
- in the dilute gas limit electrons collisions are rare, moreover electron-electron are long range (favoring forward scattering) and below  $T^*$  all these interactions are mediated by acoustic plasmons (providing partial screening)
- the above described screening, taken together with a substantial imaginary part of plasmons' velocity, justifies the plasmasonic picture
- only fermions are coupled with  $\mathbf{E}, \mathbf{B}$  fields and one then assumes their slightly shifted distribution, as it is routinely done in a dilute limit of weakly doped semiconductors, Ref.26
- in the lowest temperatures phonon-phonon collisions (in particular backward) are rare, the strongest "force" phonons receive is from the fermion drag

<sup>2</sup> note particularly long elastic scattering time for electrons in bismuth

## C Self-energy

We compute  $\Sigma_p(q, \omega)$ , the self-energy of plasmons moving through the bath of phonons. Upon its propagation, a plasmon with momentum-energy  $(q, \omega)$  emits and then reabsorbs a phonon with momentum-energy  $q', \omega_{ph}$ .  $\Sigma_p$  reads:

$$\Sigma_p(q, \omega) = \int dq' \sum_{\omega_1} \Upsilon(q, q') G_{ph}(p = q - q', \omega_1) G_{pl}(q', \omega - \omega_1) \Upsilon_0(q, q') \quad (24)$$

where integral is over all possible momenta  $q'$  of an emitted phonon. The plasmon Greens function in spectral representation:

$$G_{pl}(q, \Omega_{pl}) = \bar{A}_{pl}(q) \left( \frac{1}{\Omega_{pl} - c_{pl}q} - \frac{1}{\Omega_{pl} + c_{pl}q} \right)$$

where  $\bar{A}_{pl}(q) = \left( \frac{3\pi^2 m_r}{k_F^3} \right) \frac{c_{pl}^3 |q|}{2}$ . We perform summation over the bosonic Matsubara frequency  $\omega_1$  to get (after analytic continuation):

$$\Sigma_p(q, \omega) = \int dq' \frac{\Upsilon_0 \Upsilon \bar{A}_{pl}(q) (b(\beta c_{ph} q') + b(\beta c_{pl}(q - q')))}{(\omega + c_{ph} q' \pm c_{pl}(q - q') + i\delta)} \quad (25)$$

Results of Sec.3 allow for crucial simplifications. The coupling is resonant, which means that for each incoming plasmon with momentum  $\mathbf{q}_{pl}$  only a phonon with a momentum  $|\mathbf{q}_{ph}| = 1/2 |\mathbf{q}_{pl}|$  and  $\mathbf{q}_{ph} \parallel \mathbf{q}_{pl}$  fits the resonance condition. This introduces  $\delta(\mathbf{q}' - \mathbf{q}/2)$  and the integral over  $q'$  becomes trivial. On the resonance  $2q' = q$  and this allows to extract  $1/q$  factor in the denominator of  $\Sigma_p(q, \omega)$ . This leads to:

$$Im[\Sigma_p(q, \omega)] = \frac{2\Upsilon_0 \Upsilon \bar{A}_{pl}(q)}{q} (b(\beta c_{ph, pl} q')) \delta((\omega/q - [c_{pl} - (c_{pl} \pm c_{ph})/2])^2 + Im[c_{pl}]). \quad (26)$$

We obtain an expected result (see Fig.1) that  $Im[\Sigma_p(q, \omega)] \neq 0$  only for special values of velocities  $\omega/q = c_{pl} - (c_{pl} \pm c_{ph})/2$  and when there is a finite spectral weight along these lines. The spectral function  $Im[G_{pl}(q, \Omega_{pl})]$  is a lorentzian centered around  $\omega/q = Re[c_{pl}]$  with width  $\sim Im[c_{pl}]$ . Then  $Im[\Sigma_p(q, \omega)] \neq 0$  only when  $(c_{ph}) \leq Im[c_{pl}]$ . This is the physics presented in Fig.1, the resonance can be reached only if plasmon spectral weight is broad enough and/or velocity of phonon is close enough to velocity of plasmon.

About the momentum dependence of  $Im[\Sigma_p(q, \omega)]$ : the  $1/q$  factor cancels with momentum dependence of  $A_{pl}(q, \omega > 0) \sim q$ , so we are left with vertexes, the bare coupling has dependence  $\Upsilon_0 \sim q_{ph}^\alpha$ , see Eq.10. Then, in the lowest order, the plasmon-phonon scattering rate has the following momentum dependence:

$$\tau_p^{-1} = Im[\Sigma_p(q, \omega)] \sim q^{2\alpha} b(\beta c_{ph} q/2). \quad (27)$$

### D Three leg bubble $\Pi^{III}(\omega, \mathbf{q}, \mathbf{q}')$ : details of calculations

In order to evaluate right hand side of Eq.8 we begin with the sum over the internal Matsubara frequency of fermions. Summing over  $\omega_1$  in Eq.8 leads to three terms corresponding to the three poles:

$$\Pi^{III}(\omega, \mathbf{q}, \mathbf{q}', \Omega_2^{ph}) = \int_0^\infty d^3\mathbf{k} \left[ \frac{f(\epsilon_{\mathbf{k}-\mathbf{q}})}{(\imath\omega_n + \epsilon_{\mathbf{k}-\mathbf{q}} - \epsilon_{\mathbf{k}})(\imath(\omega_n + \Omega_2^{ph}) - \epsilon_{\mathbf{k}-\mathbf{q}} + \epsilon_{\mathbf{k}-\mathbf{p}})} + \frac{f(\epsilon_{\mathbf{k}})}{(\imath\omega_n - \epsilon_{\mathbf{k}} + \epsilon_{\mathbf{k}-\mathbf{q}})(\imath\Omega_2^{ph} - \epsilon_{\mathbf{k}-\mathbf{p}} + \epsilon_{\mathbf{k}})} + \frac{f(\epsilon_{\mathbf{k}-\mathbf{p}})}{(\imath(\omega_n - \Omega_2^{ph}) + \epsilon_{\mathbf{k}-\mathbf{q}} - \epsilon_{\mathbf{k}-\mathbf{p}})(-\imath\Omega_2^{ph} + \epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}-\mathbf{p}})} \right], \quad (28)$$

Each of these terms has a polynomial of the second order in the denominator. There are no trivial cancellations between these terms, in the following each of them will be treated separately. Since,  $V_F \mathbf{k}_F \gg k_B T^*$  we are allowed to take zero temperature limit. We make a proper shift of momentum variables to always have  $f(\epsilon_{\mathbf{k}})$  in the numerators. Then the Fermi distributions present in the numerators (Eq.28) simply restricts the limits of momentum integrals to  $\mathbf{k} = \pm \mathbf{k}_F$  at  $T = 0$ , we shall also use the fact that free-fermion dispersion is an even function of  $k$ . Then we arrive at Eq.9 in the main text.

Next we take the phonon frequency to zero ( $\Omega^{ph} \rightarrow 0$ ). This significantly simplifies the calculation and is not unreasonable as the phonon frequency is always the smallest energy scale in the problem. Nevertheless we also performed auxiliary calculations to check that for quantities of our interest, the first order corrections  $\sim \Omega^{ph}$  are indeed always one order of magnitude smaller. From qualitative analysis (see Fig.1) we see that although finite  $\Omega^{ph}$  broadens the resonance, it also increases the density of states of the plasmons involved in the process. The two effects likely compensate.

A closer insight into Eq.8 provides mathematical explanation of the resonance. If one make a shift by  $q/2$  then he/she can notice that the second and third fermionic poles gives rise to terms like  $(\epsilon_{\mathbf{k}-\mathbf{p}} - \epsilon_{\mathbf{k}-\mathbf{q}/2} + \Delta)$  (where small  $\Delta$  is related to phonons dynamics) in the denominators. In fact the one-loop master integrals, closely related to the three leg bubble  $\Pi^{III}(\omega, \mathbf{q}, \mathbf{q}')$ , is known to have a resonance feature.

In Eq.9 one needs to perform a momentum integral. Like for the case of Lindhard function ( $\Pi^{II}(\omega, \mathbf{q})$  in our notation) the momentum integral is easiest to perform in the cylindrical coordinates: if we take  $z \parallel \mathbf{q}$  then  $\int d^3\mathbf{k} \rightarrow \int dk k \sin \vartheta$  and  $(\epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}}) \rightarrow (q^2/2m - qk \cos \vartheta)$ , where  $\vartheta \angle (\mathbf{q}, \mathbf{k})$ . Firstly, one integrates over angle  $\vartheta$  and then over  $|k|$ . The analytical formulas are available only for the specific case when  $p \parallel q'$ , then  $\Pi^{III}(\omega, \mathbf{q}, \mathbf{q}') \rightarrow \Pi^{III}(\omega, q, q')$ . However we did check by numerical integration that the  $\Pi^{III}(\omega, \mathbf{q}, \mathbf{q}')$  indeed has a broad maximum at around  $\mathbf{p} \parallel \mathbf{q}'$  and is close to zero for the angles  $\angle(\mathbf{p}, \mathbf{q}) \in (\pi/3; 2\pi/3)$ . The following results is found for the three terms of Eq.9:

$$\begin{aligned} \Pi_1(q, p, w) = & \frac{q(p-q)}{3(p-2q)(q(p-q)+w)} \left( -\frac{((p-q)^2-w)^3 \log(-p^2+2pq-p-q^2+q+w)}{(p-q)^3} - \log((q-1)q-w) + 2q - \right. \\ & \frac{((p-q)^2-w)^3 \log(-2pq+(p-1)p+q^2+q-w)}{(p-q)^3} + \frac{((p-q)^2-w)^3 \log(w-(p-q)^2)}{(p-q)^3} + \log(q^2+q-w) + \\ & \log(-2pq+(p-1)p+q^2+q-w) - \frac{w(p-2q)}{q(p-q)} + \frac{((p-q)^2-w)^3 \log((p-q)^2-w)}{(p-q)^3} - \log((p-q)(p-q+1)-w) - p + \\ & \left. \frac{(q^2-w)^3 \log(q^2+q-w)}{q^3} + \frac{(q^2-w)^3 \log(-q^2+q+w)}{q^3} - \frac{(q^2-w)^3 \log(q^2-w)}{q^3} - \frac{(q^2-w)^3 \log(w-q^2)}{q^3} \right) \quad (29) \end{aligned}$$

$$\begin{aligned} \Pi_2(q, p, w) = & -\frac{q}{3(q(p+q)-w)} \left( -2p^3 \tanh^{-1}(2p+1) - 2p^3 \tanh^{-1}(1-2p) - p - \log(-p-1) + \log(1-p) - \right. \\ & \log(q^2+q-w) + \frac{(q^2-w)^3 \log(-(q^2-w)^2)}{q^3} - \frac{(q^2-w)^3 \log(-q^4+q^2(2w+1)-w^2)}{q^3} + \frac{w}{q} + \log((q-1)q-w) - q \left. \right) \quad (30) \end{aligned}$$

$$\begin{aligned}
\Pi_3(q, p, w) = & -\frac{(p-q)}{3(-pq+q^2+w)} \left( \log(p^2-2pq+p+(q-1)q+w) - \right. \\
& \frac{1}{(p-q)^3} \left[ ((p-q)^2+w)^3 \left( -\log(-p^2+2pq+p-q^2-q-w) - \log(p^2-2pq+p+(q-1)q+w) + \right. \right. \\
& \left. \left. \log(-(p-q)^2-w) + \log((p-q)^2+w) \right) + 2p^3(p-q)^3 \tanh^{-1}(2p+1) + 2p^3(p-q)^3 \tanh^{-1}(1-2p) \right] \\
& \left. - \log\left(-2pq+(p-1)p+q^2+q+w\right) + \frac{w}{p-q} + \log(p-1) - \log(p+1) - q \right) \quad (31)
\end{aligned}$$

These results are actually close analogues of the so called "*master integrals*", known from the loop expansion in the high energy physics. Like the *master integrals*, our result is also rather complicated combination of logarithms and di-logarithm functions. The di-logarithm functions are present in the more complicated cases: for the four leg bubble  $\Pi^{IV}(\omega, \mathbf{q}, \mathbf{q}')$ , for non-adiabatic case (or propagators with finite life-times) or when  $\mathbf{p}$  is not parallel to  $\mathbf{q}$ .

## References

1. A. Pourret, H. Aubin, J. Lesueur, C. A. Marrache-Kikuchi, L. Berge, L. Dumoulin, and K. Behnia. Observation of the nernst signal generated by fluctuating cooper pairs. *Nat Phys*, 2:683, 2006.
2. Olivier Cyr-Choiniere, R. Daou, Francis Laliberte, David LeBoeuf, Nicolas Doiron-Leyraud, J. Chang, J.-Q. Yan, J.-G. Cheng, J.-S. Zhou, J. B. Goode-nough, S. Pyon, T. Takayama, H. Takagi, Y. Tanaka, and Louis Taillefer. Enhancement of the nernst effect by stripe order in a high- $T_c$  superconductor. *Nature*, 458:743, 2009.
3. Jean-Philippe Brantut, Charles Grenier, Jakob Meineke, David Stadler, Sebastian Krinner, Corinna Kollath, Tilman Esslinger, and Antoine Georges. A thermoelectric heat engine with ultracold atoms. *Science*, 342(6159):713–715, 2013.
4. Lon E. Bell. Cooling, heating, generating power, and recovering waste heat with thermoelectric systems. *Science*, 321(5895):1457–1461, 2008.
5. Joseph P. Heremans. Thermoelectricity: The ugly duckling. *Nature*, 508(7496):327–328, 2014.
6. Daniel Kraemer, Bed Poudel, Hsien-Ping Feng, J. Christopher Caylor, Bo Yu, Xiao Yan, Yi Ma, Xiaowei Wang, Dezhi Wang, Andrew Muto, Kenneth McEnaney, Matteo Chiesa, Zhifeng Ren, and Gang Chen. High-performance flat-panel solar thermoelectric generators with high thermal concentration. *Nat Mater*, 10(7):532–538, 2011.
7. Bed Poudel, Qing Hao, Yi Ma, Yucheng Lan, Austin Minnich, Bo Yu, Xiao Yan, Dezhi Wang, Andrew Muto, Daryoosh Vashaee, Xiaoyuan Chen, Junming Liu, Mildred S. Dresselhaus, Gang Chen, and Zhifeng Ren. High-thermoelectric performance of nanostructured bismuth antimony telluride bulk alloys. *Science*, 320(5876):634–638, 2008.
8. Gerrit E. W. Bauer, Eiji Saitoh, and Bart J. van Wees. Spin caloritronics. *Nat Mater*, 11(5):391–399, 2012.
9. V. S. Oudovenko and G. Kotliar. Thermoelectric properties of the degenerate hubbard model. *Phys. Rev. B*, 65:075102, Jan 2002.
10. Brage Norin. Temperature and pressure dependence of the band structure in bismuth. *Physica Scripta*, 15(5-6):341, 1977.
11. N. P. Armitage, Riccardo Tediosi, F. Lévy, E. Giannini, L. Forro, and D. van der Marel. Infrared conductivity of elemental bismuth under pressure: Evidence for an avoided lifshitz-type semimetal-semiconductor transition. *Phys. Rev. Lett.*, 104(23):237401, Jun 2010.
12. Kamran Behnia, Marie-Aude Méasson, and Yakov Kopelevich. Oscillating nernst-ettingshausen effect in bismuth across the quantum limit. *Phys. Rev. Lett.*, 98:166602, Apr 2007.
13. Kamran Behnia, Luis Balicas, and Yakov Kopelevich. Signatures of electron fractionalization in ultraquantum bismuth. *Science*, 317(5845):1729–1731, 2007.
14. Huan Yang, Benot Fauqu, Liam Malone, Arlei B. Antunes, Zengwei Zhu, Ctirad Uher, and Kamran Behnia. Phase diagram of bismuth in the extreme quantum limit. *Nat. Comm.*, 1:47, 2010.
15. Yi Liu and Roland E. Allen. Electronic structure of the semimetals bi and sb. *Phys. Rev. B*, 52(3):1566–1577, Jul 1995.
16. Yuki Fuseya, Masao Ogata, and Hidetoshi Fukuyama. Transport properties and diamagnetism of dirac electrons in bismuth. *Journal of the Physical Society of Japan*, 84(1):012001, 2015.
17. P. Chudzinski and T. Giamarchi. Collective excitations and low-temperature transport properties of bismuth. *Phys. Rev. B*, 84:125105, Sep 2011.
18. David Pines and J. Robert Schrieffer. Approach to equilibrium of electrons, plasmons, and phonons in quantum and classical plasmas. *Phys. Rev.*, 125(3):804–812, Feb 1962.
19. G. D. Mahan. *Many-Particle Physics*. Physics of Solids and Liquids. Kluwer Academic/Plenum Publishers, New York, third edition, 2000.
20. L. Wendler, R. Haupt, and R. Pechstedt. Coupled plasmon-phonon excitations in quasi-one-dimensional quantum-well wires. *Phys. Rev. B*, 43(18):14669–14673, Jun 1991.
21. L. Wendler and R. Pechstedt. Dynamical screening, collective excitations, and electronphonon interaction in heterostructures and semiconductor quantum wells. general theory. *phys. stat. sol. (b)*, 138(1):197–217, 1986.
22. Riccardo Tediosi, N. P. Armitage, E. Giannini, and D. van der Marel. Charge carrier interaction with a purely electronic collective mode: Plasmarons and the infrared response of elemental bismuth. *Phys. Rev. Lett.*, 99(1):016406, Jul 2007.
23. George Whitfield and P. M. Platzman. Simultaneous strong and weak coupling in the piezoelectric polaron. *Phys. Rev. B*, 6(10):3987–3992, Nov 1972.
24. V. L. Gurevich, V. G. Skovob, and Yu. A. Firsov. Drag. *Sov. Phys.-JETP*, 13:552, 1961.
25. M. Matsuo, A. Endo, N. Hatano, H. Nakamura, R. Shirasaki, and K. Sugihara. Quantum nernst effect in a bismuth single crystal. *Phys. Rev. B*, 80(7):075313, Aug 2009.
26. T. M. Fromhold, P. N. Butcher, G. Qin, B. G. Mulimani, J. P. Oxley, and B. L. Gallagher. Phonon-drag magnetothermopower oscillations in  $\text{GaAs}/\text{As}_x\text{Ga}_{1-x}\text{As}$  heterojunctions. *Phys. Rev. B*, 48:5326–5332, Aug 1993.
27. Kamran Behnia, Marie-Aude Méasson, and Yakov Kopelevich. Nernst effect in semimetals: The effective mass and the figure of merit. *Phys. Rev. Lett.*, 98(7):076603, Feb 2007.
28. Aritra Banerjee, Benoît Fauqué, Koichi Izawa, Atsushi Miyake, Ilya Sheikin, Jacques Flouquet, Bertrand Lenoir, and Kamran Behnia. Transport anomalies across the quantum limit in semimetallic  $\text{Bi}_{0.96}\text{Sb}_{0.04}$ . *Phys. Rev. B*, 78(16):161103, Oct 2008.
29. C. Uher and W. P. Pratt. High-precision, ultralow-temperature resistivity measurements on bismuth. *Phys. Rev. Lett.*, 39(8):491–494, Aug 1977.

30. Ko Sugihara. Thermomagnetic effects in bismuth. ii. nernst-ettingshausen effect. *Journal of the Physical Society of Japan*, 27(2):362–370, 1969.
31. O. Hansen, I. Mikhail, M. Lavrenyuk, and N. Minina. Bismuth deformation potentials calculated from sdh periods under static strain and from magnetoacoustic attenuation. *Journal of Low Temperature Physics*, 95:481–496, 1994.
32. Yu. V. Sharlai and G. P. Mikitik. Oscillations of the nernst coefficient in bismuth. *Phys. Rev. B*, 83:085103, Feb 2011.
33. C. Uher and W. P. Pratt Jr. Thermopower measurements on bismuth from 9k down to 40 mk. *Journal of Physics F: Metal Physics*, 8(9):1979, 1978.
34. Y. Kuramoto. Dynamical effects of electron-hole correlation and giant quantum attenuation of ultrasound in semimetals. *Zeitschrift fr Physik B Condensed Matter*, 48:305–312, 1982. 10.1007/BF01305189.
35. K. Kajimura, H. Tokumoto, R. Inaba, and N. Mikoshiba. Anomalous behavior of giant quantum attenuation in bismuth. *Phys. Rev. B*, 12(12):5488–5501, Dec 1975.
36. Toshiyuki Yamada. Acoustomagnetolectric effect in bismuth. *Journal of the Physical Society of Japan*, 20(8):1424–1437, 1965.
37. H. Cercellier, C. Monney, F. Clerc, C. Battaglia, L. Despont, M. G. Garnier, H. Beck, P. Aebi, L. Patthey, H. Berger, and L. Forró. Evidence for an excitonic insulator phase in  $1t-tiSe_2$ . *Phys. Rev. Lett.*, 99(14):146403, Oct 2007.
38. M. Porer, U. Leierseder, J.-M. Mnard, H. Dachraoui, L. Mouchliadis, I. E. Perakis, U. Heinzmann, J. Dem-sar, K. Rossnagel, and R. Huber. Non-thermal separation of electronic and structural orders in a persisting charge density wave. *Nat. Mater.*, 13:857, Sep 2014.
39. F. Weber, R. Hott, R. Heid, K.-P. Bohnen, S. Rosenkranz, J.-P. Castellán, R. Osborn, A. H. Said, B. M. Leu, and D. Reznik. Optical phonons and the soft mode in  $2h-nbSe_2$ . *Phys. Rev. B*, 87:245111, Jun 2013.
40. B. Zenker, H. Fehske, H. Beck, C. Monney, and A. R. Bishop. Chiral charge order in  $1t-tiSe_2$ : Importance of lattice degrees of freedom. *Phys. Rev. B*, 88:075138, Aug 2013.
41. M. Jonson and G. D. Mahan. Mott’s formula for the thermopower and the wiedemann-franz law. *Phys. Rev. B*, 21:4223–4229, May 1980.